Approximating Metrics by Tree Metrics of Small Distance-Weighted Average Stretch

Mong-Jen Kao*[1], D.T. Lee†[1], and Dorothea Wagner[2]

[1]Department of Computer Science and Information Engineering, , National Taiwan University, Taipei, Taiwan., d97021@csie.ntu.edu.tw, dtlee@nchu.edu.tw
[2]Faculty of Informatics, , Karlsruhe Institute of Technology (KIT), Germany., dorothea.wagner@kit.edu

Abstract

We study the problem of how well a tree metric is able to preserve the sum of pairwise distances of an arbitrary metric. This problem is closely related to low-stretch metric embeddings and is interesting by its own flavor from the line of research proposed in the literature.

As the structure of a tree imposes great constraints on the pairwise distances, any embedding of a metric into a tree metric is known to have maximum pairwise stretch of $\Omega(\log n)$. We show, however, from the perspective of average performance, there exist tree metrics which preserve the sum of pairwise distances of the given metric up to a small constant factor, for which we also show to be no worse than twice what we can possibly expect. The approach we use to tackle this problem is more direct compared to a previous result of [4], and also leads to a provably better guarantee. Second, when the given metric is extracted from a Euclidean point set of finite dimension d, we show that there exist spanning trees of the given point set such that the sum of pairwise distances is preserved up to a constant which depends only on d. Both of our proofs are constructive. The main ingredient in our result is a special point-set decomposition which relates two seemingly-unrelated quantities.

1 Introduction

The problem of approximating a given metric by a metric which is structurally simpler has been a central issue to the theory of finite metric embedding and has been studied extensively in the past decades. A particularly simple metric of interest, which also favors from the algorithmic perspective, is a tree metric. By a tree metric we mean a metric induced by the shortest distances between pairs of points in a tree containing the given points. Generally we would require the distances in the given metric not to be underestimated in the target metric, which is crucial for most of the applications, and we would like to bound the increase of the distances, distortion, or stretch, from above. See [2, 6, 9, 12]. On the other hand, a similar and equally important problem in network design is to find a tree spanning the network, represented by a graph, that provides a good approximation to the shortest path metric defined in the graph [2, 5, 11].

Let $\mathcal{M} = (V, d)$ and $\mathcal{M}' = (V, d')$ be two metrics over the same point set V such that $d'(u, v) \geq d(u, v)$ for all $u, v \in V$. For each $u, v \in V$, let $stretch(u, v) = d'(u, v)/d(u, v)$ be the pairwise stretch, or distortion, between the pair u and v. Different notions have been suggested to quantify how well the distances of $\mathcal{M}$ are preserved in $\mathcal{M}'$, e.g.,

*This work was done when the author was with Karlsruhe Institute of Technology (KIT), as a visiting scholar under the NSC-DAAD-sponsored sandwich program (grant number NSC99-2911-I-002-055-2).

†Also with Academia Sinica, Taiwan. The author's present address is Department of Computer Science and Engineering, National Chung-Hsing University, Taichung, Taiwan.

1. Maximum pairwise stretch [15], defined by $\max_{u,v\in V} stretch(u,v)$, which is closely related to the extensively studied *Spanner* problems.

2. Average pairwise stretch [2, 11], defined by $\left(\sum_{u,v\in V} stretch(u,v)\right)/\binom{|V|}{2}$.

3. Distance-weighted average stretch [13, 16, 23], defined as

$$\frac{1}{\sum_{u,v\in V} d(u,v)} \sum_{u,v\in V} d(u,v)\cdot stretch(u,v) = \frac{\sum_{u,v\in V} d'(u,v)}{\sum_{u,v\in V} d(u,v)}.$$

   This measure makes sense in real-time scenarios when it is less desirable and more costly to raise the distances of distant pairs than that of close pairs. For example, the effect of raising the delay of a pair from 2 seconds to 10 seconds is less tolerable than raising the delay of another pair from 20 ms to 100 ms. Throughout this paper we will also refer to the sum of pairwise distances as the routing cost following the terminology used in the literature.

In this work, we address the problem of how well a tree is able to preserve the sum of pairwise distances, or, the distance-weighted average stretch, of an underlying metric. To be more precise, let $\mathcal{M} = (V, d)$ and $\mathcal{M}' = (V', d')$ be two metrics. We say that $\mathcal{M}'$ dominates $\mathcal{M}$ if $V' \supseteq V$ and for all $u, v \in V$, we have $d'(u,v) \geq d(u,v)$. We consider the following two problems.

Problem 1. Let $\mathcal{M} = (V, d)$ be a given metric and $\mathcal{D}(\mathcal{M})$ be the set of dominating tree metrics of $\mathcal{M}$. What is

$$\inf_{(V',d')\in\mathcal{D}(\mathcal{M})} \frac{\sum_{u,v\in V} d'(u,v)}{\sum_{u,v\in V} d(u,v)} \ ?$$

Problem 2. Let V be a set of points in $\mathcal{R}^d$, $|\overline{u,v}|$ be the straight-line distance between two points $u, v \in V$, $\mathcal{ST}(V)$ be the set of spanning trees of V, and $d_{\mathcal{T}}$ be the distance function of $\mathcal{T}$, for any $\mathcal{T} \in \mathcal{ST}(V)$. What is

$$\inf_{\mathcal{T}\in\mathcal{ST}(V)} \frac{\sum_{u,v\in V} d_{\mathcal{T}}(u,v)}{\sum_{u,v\in V} |\overline{u,v}|} \ ?$$

We remark on Problem 2 that, although we can consider the Euclidean metric extracted from V as we did in Problem 1, dominating tree metrics of it do not necessarily correspond to any spanning tree of V. In fact, if we apply the approaches for Problem 1 directly, the lack of balance guarantee in each partition can make the resulting pairwise distances arbitrary large.

Embedding metrics into tree metrics was introduced in the context of probabilistic embedding by Alon et al., [5]. What follows was a series of notable work. Bartal [6] considered probabilistic embeddings and proved that any metric can be probabilistically approximated by tree metrics with expected maximum distortion $O(\log^2 n)$. This result was later improved to $O(\log n \log\log n)$ [7]. Bartal also observed that any probabilistic embedding into a tree has distortion at least $\Omega(\log n)$. This gap was closed by Fakcharoenphol et al., [12], who showed that for any metric, there exists tree metrics with $O(\log n)$ distortion.

Problem 3. Given a metric $M = (V, d)$ and a weight function $w : V \times V \to \mathcal{R}^+$, find a dominating tree metric T of M such that $\sum_{u,v\in V} w_{uv}\cdot d_T(u,v) \leq \alpha \sum_{u,v\in V} w_{uv}\cdot d(u,v)$.

As Charikar et al., [10] showed by linear program duality that computing probabilistic embeddings of a given metric and Problem 3 described above are in fact dual problems, the series of work led by Bartal [6, 7, 11, 12] has provided improved approximation results for a large set of problems, including *buy-at-bulk network design*, *vehicle routing*, *metric labeling*,

group Steiner tree, *Minimum cost communication network*. Refer to [7, 10] for more detail and applications.

Kleinberg, Slivkins, and Wexler [14] initiated the study of partial embedding and scaling distortion, which can be regarded as embedding with relaxed guarantees. In a series of following work, Abraham et al., [1, 4] proved that any finite metric embeds probabilistically in a tree metric such that the distortion of $(1-\epsilon)$ portion of the pairs is bounded by $O(\log \frac{1}{\epsilon})$, for any $0 < \epsilon < 1$. They also observed a lower bound of $\Omega(\sqrt{1/\epsilon})$, which is closed by Abraham et al., in [2].

In particular, Abraham et al., [4] showed that any metric can be probabilistically embedded into a tree metric such that the ratio between the expected sum of pairwise distances is $O(\log \Phi)$, where Φ is the effective aspect ratio of given distribution. This provides an upper-bound to Problem 1 we considered. However, the guarantee they provided is loose due to the constant inherited from the guarantee on scaling distortion. See also [1, 3, 2]. Rabinovich [16] showed that it is possible to embed certain special graph metrics into real line such that distance-weighted average stretch is bounded by a constant.

On the other hand, for approximating arbitrary graph metrics by their spanning trees, a simple $\Omega(n)$ lower bound in terms of maximum stretch is known for n-cycles [17]. Alon, Karp, Peleg, and West [5] considered a distribution over spanning trees and proved an upper bound of $2^{O(\sqrt{\log n \log\log n})}$ on the expected distortion. Elkin et al., [11] showed how a spanning tree with $O(\log^2 n \log\log n)$ average stretch (over the set of edges) can be computed in polynomial time. In terms of average pairwise stretch, Abraham et al., [2] showed the existence of a spanning tree such that, for any $0 < \epsilon < 1$, the distortion of an $(1-\epsilon)$ fraction of the pairs is bounded by $O(\sqrt{1/\epsilon})$. Note that this implies an $O(1)$ average pairwise stretch.Smid [18] gave a simpler proof for this result when the metric is Euclidean.

In terms of sum of pairwise distances in graphs (routing cost), Johnson et al., [13] showed that computing the spanning tree of minimum routing cost is NP-hard. Polynomial time approximations as well as approximation schemes have been proposed by Wong [19] and Wu et al., [23]. Despite the efforts devoted, however, no general guarantees have been made on the ratio between the routing cost of the optimal spanning tree and that of the underlying graphs. Other reasonable variations have been considered as well, i.e., *sum-requirement routing trees*, *product-requirement routing trees*, and *multi-sources routing trees* [20, 22, 21].

Our Contribution In this work, we take a different approach to tackle Problem 1 directly and obtain a provably small upper-bound. Specifically, we adopt the notion of *hierarchically well-separated trees* (HSTs), introduced by Bartal [7] and Fakcharoenphol [12], and show that, for any given metric $\mathcal{M}$, there exists a 2-HST, $\mathcal{M}'$, such that the distance-weighted average stretch of $\mathcal{M}'$ is bounded by 14.24. The main ingredient of this result is a special point-set decomposition which relates two seemingly-unrelated quantities, namely, the diameter of the point set and the sum of pairwise distances between two separated subsets.

If we do not require HSTs, it is also possible to apply our technique and construct the so-called *ultra-metrics*, which is introduced by Abraham [2] and Bartal [8], with a similar stretch, 3.56. This provides a better and explicit guarantee than that provided in [4] (from ≥ 64). For the negative side, we show that there exist metrics for which no dominating tree metrics can preserve the sum of pairwise distances to a factor better than 2. This shows that our result is within twice the best one can achieve.

As a side-product, we prove the existence of spanning trees with $O(d\sqrt{d})$ distance-weighted average stretch for any point set in Euclidean space $\mathcal{R}^d$. To this end, we use our point-set cutting lemma to decompose the points recursively. In order to guarantee a constant blow-up in the diameter of the spanning tree, however, instead of allowing arbitrary cuts, we show that it is always possible to make a balanced decomposition such that the diameters of the partitioned sets stay balanced. Our result provides a good guarantee when the dimension of the given

Euclidean graph is low, which is true for most communication network. Although it is possible to apply the framework of [3, 2] to obtain a spanning tree of constant distance-weighted average stretch, the constant hidden inside is huge ($> 10^5$) that makes it practically less useful. Both of our proofs are constructive.

2 Preliminary

First we define some notation that will be used throughout this paper. Let (M, d) be a finite metric space, where M is the set of vertices and d is the distance function. Without loss of generality, we shall assume that the smallest distance defined by d is strictly more than 1. Let $X \subseteq M$ be a subset of M. The radius of X with respect to a specific element $y \in X$ is defined to be $\Delta_y(X) = \max_{z \in X} d(y, z)$. The *diameter* of X is defined to be $\Delta(X) = \max_{y \in X} \Delta_y(X)$. For any $r \geq 0$, an r-net decomposition of (M, d) is a partition of M into clusters, where each cluster, say $\mathcal{C}$, has radius at most r with respect to a certain vertex $u \in \mathcal{C}$.

Definition 1 (Hierarchical net decomposition)**.** Let (M, d) be a metric and $\delta = \lceil \log_2 \Delta(M) \rceil$. A hierarchical net decomposition of (M, d) is a sequence of $\delta + 1$ nested net decompositions $D_0, D_1, \ldots, D_\delta$ such that

- $D_\delta = \{M\}$ is the trivial partition that puts all vertices in a single cluster.
- D_i is a 2^i-net decomposition and a refinement of D_{i+1}.

A laminar family $\mathcal{F} \subseteq 2^M$ of a set M is a family of subsets of M such that for any $A, B \in \mathcal{F}$, we have either $A \subseteq B$, $B \subseteq A$, or $A \cap B = \phi$. Clearly a hierarchical net decomposition defines a laminar family and naturally corresponds to a rooted tree, for which is referred to as a hierarchically well-separated tree (HST), as follows. Each set S in the laminar family is a node in the tree, and the children of the node corresponding to S are the nodes corresponding to maximal subsets of S in the family.

The distance function on this tree is defined as follows. The links from a node S in D^i to each of its children in the tree have length equal to 2^{i-1}. This induces a distance function d_T on M, where $d_T(u, v)$ is equal to the length of the shortest path distance in T from node u to node v.

Definition 2 (Ultrametric)**.** An ultrametric M is a metric space (M, d) whose elements are the leaves of a rooted labelled tree T such that the following is met. Each node $v \in T$ is associated with a label $\ell(v) \geq 0$ such that if $u \in T$ is a descendant of v then $\ell(u) \leq \ell(v)$ and $\ell(v) = 0$ if and only if v is a leaf node. The distance between leaves $u, v \in M$ is defined as $d(u, v) = \ell(lca(u, v))$, where $lca(u, v)$ is the least common ancestor of u and v in T.

Note that, under this definition, the metric extracted from a hierarchically well-separated tree is also an ultrametric.

Definition 3 (Centripetal metric)**.** Given a metric (M, d) and a vertex $x \in M$, we define the centripetal metric (M, d_x) of (M, d) with respect to x as $d_x(u, v) = |d(u, x) - d(v, x)|$.

For any metric (X, d), we denote by $\mathcal{R}_d(X) = \sum_{u,v \in X} d(u, v)$ the sum of pairwise distances over X. Let $P, Q \subset X$ be subsets of X such that $P \cap Q = \phi$, we define $\mathcal{R}_d(P, Q) = \sum_{u \in P, v \in Q} d(u, v)$ to be the sum of pairwise distances between P and Q. The subscripts d will be omitted when there is no confusion. Clearly, $\mathcal{R}(X)$ decomposes into $\mathcal{R}(P) + \mathcal{R}(Q) + \mathcal{R}(P, Q)$ when P and Q form a partition of X.

Consider the Euclidean space of finite dimension d. A hyper-rectangle is defined to be the Cartesian product of d closed intervals, which we will denote by $[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d]$. Given a hyper-rectangle $R = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d]$, we denote by $\mathcal{L}_i(R)$ the side length

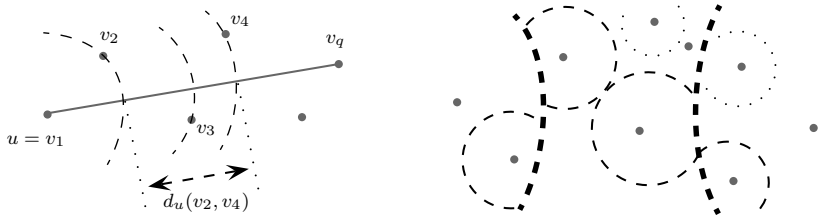


Figure 1: (a) An illustration of the centripetal metric with respect to a vertex u. (b) A hierarchical decomposition of the points.

of R along the i^{th} dimension, which is $b_i - a_i$, and $\mathcal{L}_{max}(R) = \max_{1 \le i \le d} \mathcal{L}_i(R)$. For a point set $S \in \mathcal{R}^d$, we define its bounding box, denoted by $\mathcal{B}(S)$, to be the smallest hyper-rectangle that contains S.

3 Approximating Arbitrary Metrics

Given a metric $M = (V, d)$, we describe in this section how a tree metric with small constant distance-weighted average stretch can be computed.

3.1 The Algorithm

We describe an algorithm to decompose M and define a hierarchical net decomposition. The algorithm runs in $\delta = \lceil \log_2 \Delta(V) \rceil$ iterations. Initially, we have $i = \delta$ and the trivial partition $D_\delta = \{M\}$. In each of the following iteration, we decrease the value of i by one and compute D_i from D_{i+1} as follows.

For each non-singleton cluster in D_{i+1}, say $\mathcal{P}$, we compute a 2^i-cut decomposition $\mathcal{C}(\mathcal{P})$ of $\mathcal{P}$ by repeatedly decomposing $\mathcal{P}$ by the process described below until the diameter of each clusters in the refinement falls under 2^i.

Let $\mathcal{Q}$ be a cluster in the refinement of $\mathcal{P}$ such that $\Delta(\mathcal{Q}) \ge 2^i$. We pick a vertex $u \in \mathcal{Q}$ such that $\Delta_u(\mathcal{Q}) = \Delta(\mathcal{Q})$. Then we consider the centripetal metric of $\mathcal{Q}$ with respect to u. Let $v_1, v_2, \ldots, v_q$ be the set of vertices of $\mathcal{Q}$ such that $d(u, v_1) \le d(u, v_2) \le \ldots \le d(u, v_q)$. For $1 \le i \le q-1$, we denote $\sum_{1 \le j \le i} \sum_{i < k \le q} d_u(v_j, v_k)$ by $\mathcal{RC}(i)$. Literally, $\mathcal{RC}(i)$ corresponds to the sum of pairwise distances, or, the interaction, between $\{v_1, v_2, \ldots, v_i\}$ and $\{v_{i+1}, v_{i+2}, \ldots, v_q\}$. Let p, $1 \le p < q$, be the index such that $\frac{p \cdot (q-p) \cdot \Delta(\mathcal{Q})}{\mathcal{RC}(p)}$ is minimized.We create a new cluster in the refinement of $\mathcal{P}$ containing the vertices $\{v_1, v_2, \ldots, v_p\}$ and let $\mathcal{Q} \leftarrow \mathcal{Q} \backslash \{v_1, v_2, \ldots, v_p\}$. This process is repeated until all the clusters in the refinement of $\mathcal{P}$ have diameter less than 2^i. D_i is defined to be the union of the refinements of non-singleton clusters of D_{i+1}. A high-level description of this algorithm can be found in the appendix.

3.2 Analysis

First we argue that the algorithm computes a dominating tree metric. Let T be the tree corresponding to the hierarchical net decomposition constructed by our algorithm and d_T be the distance function induced by T. For any non-singleton cluster $\mathcal{P}$ in D_i and $u, v \in \mathcal{P}$, we have $d(u, v) \le \Delta(\mathcal{P}) < 2^i$ by the definition of hierarchical net decomposition, and $d_T(u, v) \le 2 \cdot \sum_{0 \le j < i} 2^j < 2^{i+1}$ by the construction of the tree metric. Therefore, (T, d_T) is a dominating tree metric of M.

In the following, we will show that $\mathcal{R}(T) \le 4 \cdot \frac{210}{59} \cdot \mathcal{R}(M)$. To this end, we prove that, for any partition of a cluster $\mathcal{Q}$ into, say $\mathcal{Q}_1$ and $\mathcal{Q}_2$ such that $u \in \mathcal{Q}_1$, we performed in our algorithm,

we have

$$|\mathcal{Q}_1| \cdot |\mathcal{Q}_2| \cdot \Delta(\mathcal{Q}) \leq \frac{210}{59} \cdot \mathcal{R}(\mathcal{Q}_1, \mathcal{Q}_2). \tag{1}$$

Let $T[\mathcal{Q}]$, $T[\mathcal{Q}_1]$, and $T[\mathcal{Q}_2]$ denote the subtree of T corresponding to $\mathcal{Q}$, $\mathcal{Q}_1$, and $\mathcal{Q}_2$, respectively. As a consequence to (1), we have $\mathcal{R}(T_{\mathcal{Q}_1}, T_{\mathcal{Q}_2}) \leq |\mathcal{Q}_1| \cdot |\mathcal{Q}_2| \cdot 2^{i+1} \leq 4 \cdot |\mathcal{Q}_1| \cdot |\mathcal{Q}_2| \cdot \Delta(\mathcal{Q}) \leq 4 \cdot \frac{210}{59} \cdot \mathcal{R}(\mathcal{Q}_1, \mathcal{Q}_2)$. Since $\max\{|\mathcal{Q}_1|, |\mathcal{Q}_2|\} < |\mathcal{Q}|$, by an inductive argument we have $\mathcal{R}(T_{\mathcal{Q}}) = \mathcal{R}(T_{\mathcal{Q}_1}) + \mathcal{R}(T_{\mathcal{Q}_2}) + \mathcal{R}(T_{\mathcal{Q}_1}, T_{\mathcal{Q}_2}) \leq 4 \cdot \frac{210}{59} \cdot (\mathcal{R}(\mathcal{Q}_1) + \mathcal{R}(\mathcal{Q}_2) + \mathcal{R}(\mathcal{Q}_1, \mathcal{Q}_2)) = 4 \cdot \frac{210}{59} \cdot \mathcal{R}(\mathcal{Q})$. This holds for all cluster $\mathcal{Q}$, including the trivial cluster in D_δ. Therefore $\mathcal{R}(T) \leq 4 \cdot \frac{210}{59} \cdot \mathcal{R}(M)$.

It remains to prove the inequality (1). Let $\{v_1, v_2, \ldots, v_q\}$ be the set of vertices of $\mathcal{Q}$ such that $d(u, v_1) \leq d(u, v_2) \leq \ldots \leq d(u, v_q)$. Consider the following random distribution defined over $\beta \in \left\{ \lceil \frac{q}{4} \rceil, \lceil \frac{q}{4} \rceil + 1, \ldots, \left\lfloor \frac{3q}{4} \right\rfloor \right\}$.

$$Pr[\beta = i] = \frac{\mathcal{RC}(i)}{\sum_{\frac{q}{4} \leq i \leq \frac{3q}{4}} \mathcal{RC}(i)}$$

Let us first derive a lower bound on $\sum_{\frac{q}{4} \leq i \leq \frac{3q}{4}} \mathcal{RC}(i)$, which is the total amount of interaction when cutting at the central $\frac{q}{2}$ intervals. Due to space limit, preliminary material as well as proofs to the following lemmas are moved to the appendix for further reference.

Lemma 1. We have

$$\sum_{\frac{q}{4} \leq i \leq \frac{3}{4}q} \mathcal{RC}(i) \geq \left(\frac{3}{32} q^3 + \frac{q}{2} \cdot \sum_{\frac{q}{6} q \leq i \leq \frac{q}{4}} i \right) \cdot \sum_{\frac{q}{3} \leq k \leq \frac{2q}{3}} \ell_k$$

The following lemma proves the existence of a good cut and (1).

Lemma 2. We have

$$\min \left\{ E\left[\frac{\beta \cdot (q-\beta) \cdot \Delta(\mathcal{Q})}{\mathcal{RC}(\beta)} \right], \min_{1 \leq \gamma \leq \frac{q}{3}} \left\{ \frac{\gamma \cdot (q-\gamma) \cdot \Delta(\mathcal{Q})}{\mathcal{RC}(\gamma)}, \frac{\gamma \cdot (q-\gamma) \cdot \Delta(\mathcal{Q})}{\mathcal{RC}(q-\gamma)} \right\} \right\} \leq \frac{210}{59}.$$

As a side-product, we have the following lemma, which states the existence of good cuts for any given point set and the correctness of inequality (1).

Lemma 3 (1-Dimensional Point Set Cutting Lemma)**.** Given a set of real numbers $A = \{a_1, a_2, \ldots, a_n\}$, $a_1 \leq a_2 \leq \ldots \leq a_n$, there exists a cutting point $z \in R$ with $a_1 < z < a_n$ such that the following holds.

$$L_A(z) \cdot (n - L_A(z)) \cdot \Delta \leq \delta_0 \cdot \sum_{1 \leq i \leq L_A(z)} \sum_{L_A(z) < j \leq n} (a_j - a_i),$$

where $L_A(z) = |\{a \in A : a < z\}|$ is the number of elements in A that are smaller than z, $\Delta = a_n - a_1$ is the diameter of A, and $\delta_0 \leq \frac{210}{59}$ is a constant.

3.3 Lower Bound

In the following, we derive a lower bound to Problem 1 we considered throughout this section. This is done by linking the basic structure of any optimal dominating tree metric to our point set cutting lemma, followed by deriving an upper bound to the performance of any cut.

Let $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ be a set of numbers, where $a_i = i$ for all $1 \leq i \leq n$, and $(\mathcal{A}, d)$ be the corresponding metric extracted from $\mathcal{A}$. Let (T, d_T) be an optimal ultra-metric embedding of $\mathcal{A}$ in terms of distance-weighted average stretch. Without loss of generality, we can assume that T is a binary tree. Otherwise, we can always create dummy nodes to make T binary without changing its sum of pairwise distances. The following lemma characterizes the structure of T.

Lemma 4. Let T_L and T_R be the left-subtree and the right-subtree of T such that $a_1 \in T_L$. Then, there exists an integer k, $1 \leq k < n$, such that T_L is an ultra-metric containing $\{a_1, a_2, \ldots, a_k\}$ and T_R is an ultra-metric containing $\mathcal{A}\backslash\{a_1, a_2, \ldots, a_k\}$.

Therefore, to obtain a lower bound on the distance-weighted average stretch of any dominating tree metric of $\mathcal{A}$, it suffices to consider the quality of the best cut we can possibly achieve on $\mathcal{A}$.

Lemma 5. Let δ_0 be a constant such that our point set cutting lemma holds, then $\delta_0 \geq 2$.

By Lemma 4 and Lemma 5, we obtain the following bound as claimed.

Corollary 6. Let $\mathcal{M} = (V, d)$ be a given metric and $\mathcal{D}(\mathcal{M})$ be the set of dominating tree metrics of $\mathcal{M}$. Then

$$\inf_{(V',d')\in\mathcal{D}(\mathcal{M})} \frac{\sum_{u,v\in V} d'(u,v)}{\sum_{u,v\in V} d(u,v)} \geq 2.$$

4 Approximating Euclidean Metrics by Their Spanning Trees

In this section, we show how a spanning tree of small constant distance-weighted average stretch for a Euclidean graph can be computed in polynomial time. The basic idea is to extend the previous point-set decomposition. In order to guarantee a constant blow-up in the diameter of the resulting spanning tree, we cannot allow the cut to be made at arbitrary positions. Instead, we restrict each cut to be made within the central $(1-2\alpha)$ portion along the longest side of its bounding box, where α is a constant chosen to be $\frac{1}{4}$. This guarantees a balanced partition, an exponentially decreasing size of the bounding boxes, and a constant blow-up of the diameter of the resulting spanning tree. This is crucial in the analysis, as we need a tight diameter in order to provide a good upper-bound on the interaction between pairs separated by our cuts. On the other hand, we also have to guarantee the existence of good cuts in the central $(1-2\alpha)$ portion so that the overall interaction stays bounded.

Given a set of points $\mathcal{P}$ in the Euclidean space $\mathcal{R}^d$ of finite dimension, our algorithm recursively computes a rooted tree $\mathcal{T}$ with root r as follows. Let $\mathcal{B}(\mathcal{P})$ be the bounding box of $\mathcal{P}$, and k be the index of dimension such that $\mathcal{L}_k(\mathcal{B}(\mathcal{P})) = \mathcal{L}_{max}(\mathcal{B}(\mathcal{P}))$. We consider the projection of the points to the k^{th}-axis, and let $a_1, a_2, \ldots, a_n$, $a_1 \leq a_2 \leq \ldots \leq a_n$, be the corresponding coordinates. We apply our linear time algorithm[1] to compute a decomposition for which the cut is restricted to be made inside the central $(1-2\alpha)$ portion, $[\alpha \cdot (a_1 + a_n), (1-\alpha) \cdot (a_1 + a_n)]$. See also Fig. 2 (a). Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be the corresponding partitioned subsets of points. We compute recursively the two rooted trees for $\mathcal{P}_1$ and $\mathcal{P}_2$, denoted by $\mathcal{T}_1$ with root r_1 and $\mathcal{T}_2$ with root r_2. The tree $\mathcal{T}$ is constructed by joining r_1 and r_2, and the root of $\mathcal{T}$ is chosen to be r_1. A high-level description of our algorithm is provided in the appendix.

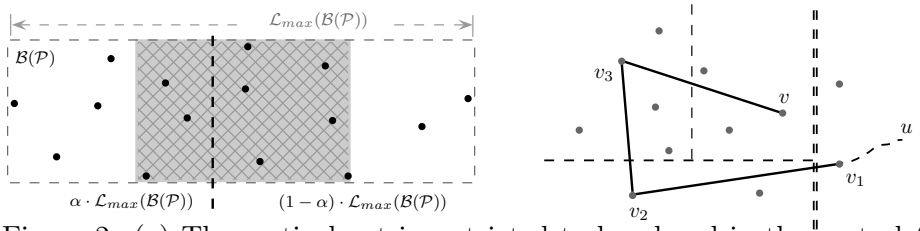


Figure 2: (a) The vertical cut is restricted to be placed in the central $(1-2\alpha)$ portion along the longest side of the bounding box. (b) A possible decomposition and the $u-v$ path in the resulting tree.

[1]This algorithm is moved to § A.4 for further reference due to space limit.

In the following lemma, we show that, in exchange of certain penalty in the performance factor that is inverse proportional to the length of the interval to which the cut is restricted, we can always guarantee a good and balanced decomposition.

Lemma 7 (Constrained Point Set Cutting Lemma)**.** Given a set of real numbers $A = \{a_1, a_2, \ldots, a_n\}$, $a_1 \leq a_2 \leq \ldots \leq a_n$ and an interval $\mathcal{I} = [\ell, r]$ such that $\mathcal{I} \subseteq [a_1, a_n]$, there exists a cutting point $z \in \mathcal{I}$ such that the following holds.

$$L_A(z) \cdot (n - L_A(z)) \cdot |\mathcal{I}| \leq \delta_0 \cdot \sum_{1 \leq i \leq L_A(z)} \sum_{L_A(z) < j \leq n} (a_j - a_i),$$

where $L_A(z) = |\{a \in A : a < z\}|$ is the number of elements in A that are smaller than z and $\delta_0 \leq \frac{210}{59}$ is a constant.

In the following, we state the theorem and leave the rest detail in the appendix for further reference.

Theorem 8. Given a set of points $\mathcal{P}$ in $\mathcal{R}^d$, we can compute in polynomial time a spanning tree $\mathcal{T}$ of $\mathcal{P}$ such that the distance-weighted average stretch of $\mathcal{T}$ with respect to $\mathcal{P}$ is at most $16\delta_0 \cdot d\sqrt{d}$, where $\delta_0 \leq \frac{210}{59}$ is the constant in our point set cutting lemma.

5 Discussion and Open Problems

We conclude with some remarks and conjectures. In this work, we provided both an upper bound and a lower bound to Problem 1. We conjecture the lower bound of two we provided to be tight. On the other hand, we also conjecture that similar result holds for approximating arbitrary graph metrics by their spanning trees. However, as it seems not promising to guarantee the quality of the best cut for arbitrarily small restricted intervals, none of known graph decomposition techniques helps and either more powerful decomposition schemes or new techniques are expected.

References

[1] I. Abraham, Y. Bartal, T-H. Chan, K. Dhamdhere, A. Gupta, J. Kleinberg, O. Neiman, and A. Slivkins. Metric embeddings with relaxed guarantees. In *FOCS'05*, pages 83–100, Washington, DC, USA, 2005.

[2] I. Abraham, Y. Bartal, and O. Neiman. Embedding metrics into ultrametrics and graphs into spanning trees with constant average distortion. In *SODA'07*, pages 502–511, Philadelphia, PA, USA, 2007.

[3] Ittai Abraham, Yair Bartal, and Ofer Neiman. On embedding of finite metric spaces into hilbert space. manuscript, 2005.

[4] Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. In *STOC '06*, pages 271–286, New York, NY, USA, 2006. ACM.

[5] N. Alon, R. Karp, D.d Peleg, and D. West. A graph-theoretic game and its application to the k-server problem. *SIAM J. Comput.*, 24:78–100, February 1995.

[6] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *FOCS'96*, pages 184–, Washington, DC, USA, 1996.

[7] Yair Bartal. On approximating arbitrary metrices by tree metrics. In *STOC'98*, pages 161–168, New York, NY, USA, 1998. ACM.

[8] Yair Bartal. Graph decomposition lemmas and their role in metric embedding methods. In *ESA'04*, pages 89–97, 2004.

[9] Yair Bartal, Nathan Linial, Manor Mendel, and Assaf Naor. On metric ramsey-type phenomena. In *STOC'03*, pages 463–472, New York, NY, USA, 2003. ACM.

[10] M. Charikar, C. Chekuri, A. Goel, S. Guha, and S. Plotkin. Approximating a finite metric by a small number of tree metrics. In *FOCS'98*, pages 379–, 1998.

[11] M. Elkin, Y. Emek, D. Spielman, and S.-H. Teng. Lower-stretch spanning trees. In *STOC'05*, pages 494–503, New York, NY, USA, 2005. ACM.

[12] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *STOC'03*, pages 448–455, New York, NY, 2003.

[13] D. S. Johnson, J. K. Lenstra, and A. H. G. Rinnooy Kan. The complexity of the network design problem. *Networks*, 8:279–285, 1978.

[14] Jon Kleinberg, Aleksandrs Slivkins, and Tom Wexler. Triangulation and embedding using small sets of beacons. *J. ACM*, 56:32:1–32:37, September 2009.

[15] Giri Narasimhan and Michiel Smid. *Geometric Spanner Networks*. Cambridge University Press, New York, NY, USA, 2007.

[16] Yuri Rabinovich. On average distortion of embedding metrics into the line. In *STOC'03*, pages 456–462, 2003.

[17] Yuri Rabinovich and Ran Raz. Lower bounds on the distortion of embedding finite metric spaces in graphs. *Discrete & Computational Geometry*, 19, 1996.

[18] Michiel Smid. Spanning trees with o(1) average stretch factor. manuscript, 2009.

[19] Richard Wong. Worst-case analysis of network design problem heuristics. *SIAM. J. Alg. Disc. Meth.*, 1, 1980.

[20] B.-Y. Wu. Approximation algorithms for the optimal p-source communication spanning tree. *Discrete Appl. Math.*, 143:31–42, September 2004.

[21] B.-Y. Wu, K.-M. Chao, and C.-Y. Tang. Light graphs with small routing cost. *Networks*, 39:2002.

[22] B.-Y. Wu, K.-M. Chao, and C.-Y. Tang. A polynomial time approximation scheme for optimal product-requirement communication spanning trees. *J. Algorithms*, 36:182–204, August 2000.

[23] B.-Y. Wu, G. Lancia, V. Bafna, K.-M. Chao, R. Ravi, and C.-Y. Tang. A polynomial-time approximation scheme for minimum routing cost spanning trees. *SIAM J. Comput.*, 29:761–778, December 1999.

A Approximating Arbitrary Metrics

A.1 The Algorithm

ALGORITHM *Hierarchical-Net-Decomposition*(V, d)

$D_\delta \leftarrow \{V\}$, $i \leftarrow \delta - 1$.
while $i \geq 0$ and D_{i+1} has non-singleton clusters **do**
  for all non-singleton cluster $\mathcal{P}$ in D_{i+1} **do**
    $\mathcal{C}(\mathcal{P}) \leftarrow \{\phi\}$, $\mathcal{S} \leftarrow \{\mathcal{P}\}$.
    while $\mathcal{S} \neq \phi$ **do**
      Let $\mathcal{Q}$ be an arbitrary cluster in $\mathcal{S}$.
      if $\Delta(\mathcal{Q}) < 2^i$ **then**
        Add $\mathcal{Q}$ to $\mathcal{C}(\mathcal{P})$ and remove $\mathcal{Q}$ from $\mathcal{S}$.
      else
        Let $u \in \mathcal{Q}$ be a vertex such that $\Delta_u(\mathcal{Q}) = \Delta(\mathcal{Q})$.
        Let $v_1, v_2, \ldots, v_q$ be the set of vertices in $\mathcal{Q}$ such that $d(u, v_1) \leq d(u, v_2) \leq \ldots \leq d(u, v_q)$.
        Let p, $1 \leq p < q$, be the index such that $\frac{p \cdot (q-p) \cdot \Delta(\mathcal{Q})}{\mathcal{RC}(p)}$ is minimized.
        Let $\mathcal{Q}' \leftarrow \{v_1, v_2, \ldots, v_p\}$, $\mathcal{S} \leftarrow \mathcal{S} \cup \{\mathcal{Q}'\}$, and $\mathcal{Q} \leftarrow \mathcal{Q} \backslash \mathcal{Q}'$.
      end if
    end while
    Let $\mathcal{C}(\mathcal{P})$ be the refinement clusters of $\mathcal{P}$ in D_i.
  end for
  $i \leftarrow i - 1$.
end while
Return the tree metric corresponding to $D_0, D_1, \ldots, D_\delta$.

Figure 3: A high-level description of the algorithm.

A.2 Analysis

Lemma 1. *We have*

$$\sum_{\frac{q}{4} \leq i \leq \frac{3}{4}q} \mathcal{RC}(i) \geq \left(\frac{3}{32} q^3 + \frac{q}{2} \cdot \sum_{\frac{q}{6} q \leq i \leq \frac{q}{4}} i \right) \cdot \sum_{\frac{q}{3} \leq k \leq \frac{2q}{3}} \ell_k$$

Before proving Lemma 1, let us derive a lower bound on the overall interaction $\sum_{1 \leq i < q} \mathcal{RC}(i)$. Recall that, $\mathcal{RC}(i) = \sum_{1 \leq j < i} \sum_{i < j \leq q} d_u(v_j, v_k)$ and $d_u(v_j, v_k) = |d(u, v_j) - d(u, v_k)|$. For convenience, we will denote by ℓ_k the quantity $d_u(v_k, v_{k+1})$, which is exactly $d(u, v_{k+1}) - d(u, v_k)$, for each $1 \leq k < q$.

First, observe that, for each j, k with $1 < j < k < q$, we have exactly $(k - j)$ duplications of the item $d_u(v_j, v_k)$ in the summation $\sum_{1 \leq i < q} \mathcal{RC}(i)$, i.e., it appears exactly once in $\mathcal{RC}(i)$ for each $j \leq i < k$. Therefore, after re-arranging the items we have

$$\sum_{1 \leq i < q} \mathcal{RC}(i) = \sum_{1 \leq k < q} k \cdot \sum_{1 \leq i \leq q-k} d_u(v_i, v_{i+k}).$$

Let $f(q) = \frac{q}{2} \sum_{1 \leq i \leq \frac{q}{2}} d_u(v_i, v_{i+\frac{q}{2}})$ if q is even and $f(q) = 0$ otherwise. Then

$$\begin{aligned}
&\sum_{1 \leq k < q} k \cdot \sum_{1 \leq i \leq q-k} d_u(v_i, v_{i+k}) \\
&= \sum_{1 \leq k < \frac{q}{2}} k \cdot \sum_{1 \leq i \leq q-k} d_u(v_i, v_{i+k}) + \sum_{\frac{q}{2} < k < q} k \cdot \sum_{1 \leq i \leq q-k} d_u(v_i, v_{i+k}) + f(q) \\
&= \sum_{1 \leq k < \frac{q}{2}} k \cdot \sum_{1 \leq i \leq q-k} d_u(v_i, v_{i+k}) + \sum_{1 \leq k < \frac{q}{2}} (q - k) \sum_{1 \leq i \leq k} d_u(v_i, v_{i+q-k}) + f(q),
\end{aligned}$$

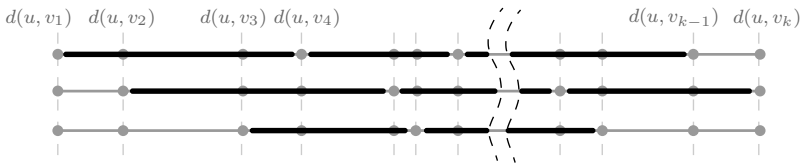


Figure 4: Alignment of the intervals when $k = 3$. The first group starts with $d(u, v_1)$ while the second and the third start with $d(u, v_2)$ and $d(u, v_3)$, respectively.

where in the last inequality we substitute the variable k by $q - k$. By re-organizing and aligning the items from the above summation (see also Fig. 4), we have the following lemma.

Lemma 9. For $1 \leq k \leq \left\lfloor \frac{q}{2} \right\rfloor$, we have

$$\sum_{1 \leq i \leq q-k} d_u(v_i, v_{i+k}) = k \cdot \Delta(\mathcal{Q}) - \sum_{1 \leq i < k} (k - i) \cdot (\ell_i + \ell_{q-i}) = \sum_{1 \leq i \leq k} d_u(v_i, v_{i+q-k})$$

Proof of Lemma 9. We prove the first half of this lemma, $\sum_{1 \leq i \leq q-k} d_v(v_{i+k}, v_i) = k \cdot \Delta(\mathcal{Q}) - \sum_{1 \leq i < k} (k - i) \cdot (\ell_i + \ell_{q-i})$. The second half, $\sum_{1 \leq i \leq k} d_v(v_{i+q-k}, v_i) = k \cdot \Delta(\mathcal{Q}) - \sum_{1 \leq i < k} (k - i) \cdot (\ell_i + \ell_{q-i})$, follows by a similar argument. Consider the alignments of the set of intervals which spans exactly k consecutive elements, that is, intervals $[d(u, v_i), d(u, v_{i+k})]$, for $1 \leq k \leq \left\lfloor \frac{q}{2} \right\rfloor$. We have exactly k alignments, each starting with $\mathcal{I}_i$ for $1 \leq i \leq k$. See also Fig. 4. This sums up to $k \cdot \Delta(\mathcal{Q})$, except for exactly $k - i$ times over-count of ℓ_i and ℓ_{q-i}. □ □

We provide in the following lemma an overall estimate to the overall interaction, $\sum_{1 \leq i < q} \mathcal{RC}(i)$.

Lemma 10.

$$\sum_{1 \leq i < q} \mathcal{RC}(i) \geq \sum_{1 \leq k < \frac{q}{2}} q \cdot \sum_{\frac{q}{2} - k < i < \frac{q}{2}} i \cdot (\ell_k + \ell_{q-k}) + g(q),$$

where $g(q) = q \cdot \sum_{1 \leq i < \frac{q}{2}} i \cdot \ell_{\frac{q}{2}}$ if q is even and $g(q) = 0$ otherwise.

Proof of Lemma 10. By the above discussion and Lemma 9, we have

$$\begin{aligned}
&\sum_{1 \leq i < q} \mathcal{RC}(i) \\
&= \sum_{1 \leq k < \frac{q}{2}} k \cdot \sum_{1 \leq i \leq q-k} d_u(v_i, v_{i+k}) + \sum_{1 \leq k < \frac{q}{2}} (q - k) \sum_{1 \leq i \leq k} d_u(v_i, v_{i+p-k}) + f(q) \\
&= \sum_{1 \leq k \leq \frac{q}{2}} q \cdot \left(k \cdot \Delta(\mathcal{Q}) - \sum_{1 \leq i < k} (k - i)(\ell_i + \ell_{q-i}) \right)
\end{aligned}$$

For $1 \leq i < \frac{q}{2}$, the coefficient of ℓ_i and ℓ_{q-i} in the above summation is $q \cdot \sum_{i < k < \frac{q}{2}} (k - i)$, which equals $q \cdot \sum_{1 \leq k < \frac{q}{2} - i} k$ by substituting the variable k by $k - i$. Therefore, we have

$$\sum_{1 \leq i < q} \mathcal{RC}(i) \geq \sum_{1 \leq k < \frac{q}{2}} q \cdot k \cdot \Delta(\mathcal{Q}) - \sum_{1 \leq k < \frac{q}{2}} q \cdot \sum_{1 \leq i < \frac{q}{2} - k} i \cdot (\ell_k + \ell_{q-k}).$$

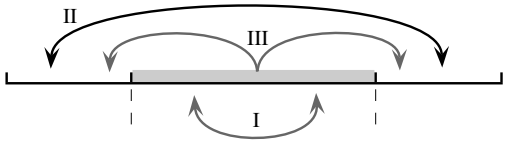
Since $\Delta(\mathcal{Q}) = \sum_{1 \leq i < q} \ell_i$, by further expanding $\Delta(\mathcal{Q})$, we obtain

$$\sum_{1 \leq i < q} \mathcal{RC}(i) \geq \sum_{1 \leq k < \frac{q}{2}} q \cdot \sum_{\frac{q}{2} - k < i < \frac{q}{2}} i \cdot (\ell_k + \ell_{q-k}) + g(q).$$

□ □

Now we are ready to prove Lemma 1.

Proof of Lemma 1. We divide the total interaction to be lower-bounded, $\sum_{\frac{q}{4} \leq i \leq \frac{3}{4}q} \mathcal{RC}(i)$, into three parts which we discuss below.



I. the interaction between points from $\left\{v_{\lceil \frac{q}{4} \rceil}, v_{\lceil \frac{q}{4} \rceil+1}, \ldots, v_{\lfloor \frac{3q}{4} \rfloor}\right\}$.

The situation is equivalent to computing the overall interaction for a point set of $\frac{q}{2}$ points. By Lemma 10 with index replacement, the interaction is lower-bounded by $\sum_{1\le k<\frac{q}{4}} \frac{q}{2} \cdot \sum_{\frac{q}{4}-k<i<\frac{q}{4}} i \cdot (\ell_{\frac{q}{4}+k}+\ell_{\frac{3q}{4}-k})+g'(q)$, where $g'(q) = \frac{q}{2}\cdot\sum_{1\le i<\frac{q}{4}} i\cdot\ell_{\frac{q}{2}}$ if $\frac{q}{2}$ is even and $g'(q)=0$ otherwise. Dropping the items corresponding to $k < \frac{q}{12}$ from the first summation, we obtain $\frac{q}{2}\cdot\sum_{\frac{q}{6}q\le i\le\frac{q}{4}} i\cdot\sum_{\frac{q}{3}\le k\le\frac{2q}{3}} \ell_k$.

For the remaining two cases, we consider the number of times each of the items from $\sum_{\frac{q}{3}\le k\le\frac{2q}{3}} \ell_k$ contributes to $\sum_{\frac{q}{4}\le i\le\frac{3q}{4}} \mathcal{RC}(i)$.

II. the interaction between $\left\{v_1, v_2, \ldots, v_{\lceil \frac{q}{4} \rceil}\right\}$ and $\left\{v_{\lfloor \frac{3q}{4} \rfloor}, v_{\lfloor \frac{3q}{4} \rfloor+1}, \ldots, v_q\right\}$.

For each j,k such that $1 \le j \le \frac{q}{4}$, $\frac{3q}{4} \le k < q$, the pair $d_u(v_j, v_k)$ contributes exactly once to the term $\mathcal{RC}(i)$ for each i with $\frac{q}{4} \le i \le \frac{3q}{4}$. There are $\frac{1}{16}q^2$ such pairs, while there are $\frac{q}{2}$ different terms in the final summation $\sum_{\frac{q}{4}\le i\le\frac{3}{4}q} \mathcal{RC}(i)$. Therefore, we obtain a lower bound of $\frac{1}{32}q^3\cdot\sum_{\frac{q}{3}\le k\le\frac{2q}{3}} \ell_k$ for this part.

III. the interaction between $\left\{v_{\lceil \frac{q}{4} \rceil}, v_{\lceil \frac{q}{4} \rceil+1}, \ldots, v_{\lfloor \frac{3q}{4} \rfloor}\right\}$ and other points.

For any specific interval ℓ_p with $\frac{q}{4} \le p \le \frac{3q}{4}$, we consider the number of pairs between $\left\{v_{\lceil \frac{q}{4} \rceil}, v_{\lceil \frac{q}{4} \rceil+1}, \ldots, v_{\lfloor \frac{3q}{4} \rfloor}\right\}$ and other points that contain this specific interval ℓ_p. There are $p-\frac{q}{4}$ points, $\left\{v_{\lceil \frac{q}{4} \rceil}, v_{\lceil \frac{q}{4} \rceil+1}, \ldots, v_p\right\}$, which lie to the left of v_p and form pairs with points from $\left\{v_{\lfloor \frac{3q}{4} \rfloor}, v_{\lfloor \frac{3q}{4} \rfloor+1}, \ldots, v_q\right\}$ that contain ℓ_p. Similarly, the $\frac{3q}{4} - p$ points that lie to the right of v_p also form pairs with points from $\left\{v_1, v_2, \ldots, v_{\lceil \frac{q}{4} \rceil}\right\}$ that contain ℓ_p. Therefore there are $\frac{q}{4}\cdot\left(p-\frac{q}{4}+\frac{3q}{4}-p\right) = \frac{q}{4}\cdot\frac{q}{2}$ such pairs. This is true for all $\mathcal{RC}(i)$ with $\frac{q}{4} \le i \le \frac{3q}{4}$. Therefore ℓ_p contributes $\frac{q}{4}\cdot\frac{q}{2}\cdot\frac{q}{2}$ times in the summation and we obtain a lower bound of $\frac{1}{16}q^3\cdot\sum_{\frac{q}{3}\le k\le\frac{2q}{3}} \ell_k$.

Summing up the bounds we obtained in the three parts and we have this lemma. □ □

Lemma 2. *We have*

$$\min\left\{E\left[\frac{\beta\cdot(q-\beta)\cdot\Delta(\mathcal{Q})}{\mathcal{RC}(\beta)}\right], \min_{1\le\gamma\le\frac{q}{3}}\left\{\frac{\gamma\cdot(q-\gamma)\cdot\Delta(\mathcal{Q})}{\mathcal{RC}(\gamma)}, \frac{\gamma\cdot(q-\gamma)\cdot\Delta(\mathcal{Q})}{\mathcal{RC}(q-\gamma)}\right\}\right\} \le \frac{210}{59}.$$

Proof of Lemma 2. This lemma holds trivially when $q \le 3$. For $q \ge 4$, by the definition of expected values, we have

$$E\left[\frac{\beta\cdot(q-\beta)\cdot\Delta(\mathcal{Q})}{\mathcal{RC}(\beta)}\right] = \sum_{\frac{q}{4}\le i\le\frac{3q}{4}} Pr[\beta=i]\cdot\frac{\beta\cdot(q-\beta)\cdot\Delta(\mathcal{Q})}{\mathcal{RC}(\beta)} = \frac{\sum_{\frac{q}{4}\le i\le\frac{3q}{4}} i\cdot(q-i)\cdot\Delta(\mathcal{Q})}{\sum_{\frac{q}{4}\le i\le\frac{3q}{4}}\mathcal{RC}(i)}.$$

First we have

$$\sum_{\frac{q}{4}\le i\le\frac{3q}{4}} i(q-i)\cdot\Delta(\mathcal{Q}) = \left(q\cdot\sum_{\frac{q}{4}\le i\le\frac{3q}{4}} i - \sum_{\frac{q}{4}\le i\le\frac{3q}{4}} i^2\right)\cdot\Delta(\mathcal{Q}) \le \frac{11}{96}q^3\Delta(\mathcal{Q}).$$

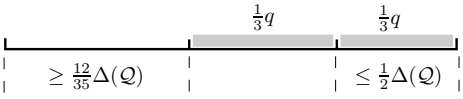
Depending on whether or not $\sum_{\frac{q}{3}\le k\le\frac{2q}{3}} \ell_i \ge \frac{11}{35}\Delta(\mathcal{Q})$, we distinguish between two cases.

If $\sum_{\frac{q}{3}\le k\le\frac{2q}{3}} \ell_i \ge \frac{11}{35}\Delta(\mathcal{Q})$, then, by Lemma 1, we have

$$\sum_{\frac{q}{4}\le i\le\frac{3q}{4}} \mathcal{RC}(i) \ge \sum_{\frac{q}{3}\le k\le\frac{2q}{3}} \ell_k\cdot\left(\frac{3}{32}q^3 + \frac{q}{2}\cdot\sum_{\frac{q}{6}\le i\le\frac{q}{4}} i\right) \ge \frac{11}{35}\Delta(\mathcal{Q})\cdot\frac{59}{96\cdot 6}q^3,$$

$$\text{and} \quad E\left[\frac{\beta\cdot(q-\beta)\cdot\Delta(\mathcal{Q})}{\mathcal{RC}(\beta)}\right] \le \frac{11}{96}q^3\Delta(\mathcal{Q})/\left(\frac{11}{35}\Delta(\mathcal{Q})\cdot\frac{59}{96\cdot 6}q^3\right) \le \frac{210}{59}.$$

On the other hand, if $\sum_{1\le i\le \frac{q}{3}}(\ell_i+\ell_{q-i}) \ge \frac{11}{35}\Delta(\mathcal{Q})$, then we have either $\sum_{1\le i\le\frac{q}{3}}\ell_i \ge \frac{12}{35}\Delta(\mathcal{Q})$, or $\sum_{1\le i\le\frac{q}{3}}\ell_{q-i} \ge \frac{12}{35}\Delta(\mathcal{Q})$. Without loss of generality, assume that $\sum_{1\le i\le\frac{q}{3}}\ell_i \ge \sum_{1\le i\le\frac{q}{3}}\ell_{q-i} \ge \frac{12}{35}\Delta(\mathcal{Q})$.



In this case, we have $\sum_{1\le i\le\frac{q}{3}}\ell_i + \sum_{\frac{q}{3}<i<\frac{2q}{3}}\ell_i \ge \sum_{\frac{2q}{3}\le i<q}\ell_i$. Therefore $\sum_{\frac{2q}{3}\le i<q}\ell_i \le \frac{\Delta(\mathcal{Q})}{2}$. Let p be the smallest integer such that $\ell_p > 0$. Counting the interaction between $\{v_1, v_2, \ldots, v_p\}$ and $\{v_{p+1}, v_{p+2}, \ldots, v_q\}$, we have $\mathcal{RC}(p) \ge p\cdot\frac{q}{3}\cdot\frac{12}{35}\Delta(\mathcal{Q}) + p\cdot\frac{q}{3}\cdot\frac{1}{2}\Delta(\mathcal{Q})$. Therefore,

$$\frac{p\cdot(q-p)\cdot\Delta(\mathcal{Q})}{\mathcal{RC}(p)} \le \frac{p\cdot q\cdot\Delta(\mathcal{Q})}{p\cdot q\cdot\Delta(\mathcal{Q})\cdot\left(\frac{1}{3}\cdot\frac{12}{35}+\frac{1}{3}\cdot\frac{1}{2}\right)} = \frac{210}{59}.$$

The argument for the case $\sum_{1\le i\le\frac{q}{3}}\ell_{q-i} \ge \sum_{1\le i\le\frac{q}{3}}\ell_i$ is analogous. This proves the lemma. □ □

A.3 Lower Bound

Let $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ be a set of numbers, where $a_i = i$ for all $1 \le i \le n$, and $(\mathcal{A}, d)$ be the corresponding metric extracted from $\mathcal{A}$. Let (T, d_T) be an optimal ultra-metric embedding of $\mathcal{A}$ in terms of distance-weighted average stretch. Without loss of generality, we can assume that T is a binary tree. Otherwise, we can always create dummy nodes to make T binary without changing its sum of pairwise distances. The following lemma characterizes the structure of T.

Lemma 4. *Let T_L and T_R be the left-subtree and the right-subtree of T such that $a_1 \in T_L$. Then, there exists an integer k, $1 \le k < n$, such that T_L is an ultra-metric containing $\{a_1, a_2, \ldots, a_k\}$ and T_R is an ultra-metric containing $\mathcal{A}\backslash\{a_1, a_2, \ldots, a_k\}$.*

Proof of Lemma 4. If not, let ℓ be the number of leaves in T_L, and denote by φ the permutation on $\{1, 2, \ldots, n\}$ such that T_L is an ultra-metric containing $\{a_{\varphi(1)}, a_{\varphi(2)}, \ldots, a_{\varphi(\ell)}\}$, where $a_{\varphi(1)} < a_{\varphi(2)} < \ldots < a_{\varphi(\ell)}$, and T_R is an ultra-metric containing $\{a_{\varphi(\ell+1)}, a_{\varphi(\ell+2)}, \ldots, a_{\varphi(n)}\}$, where $a_{\varphi(\ell+1)} < a_{\varphi(\ell+2)} < \ldots < a_{\varphi(n)}$. Note that by our assumption, $a_{\varphi(\ell)} > a_{\varphi(\ell+1)}$.

Construct a new ultra-metric $\mathcal{T}_0$ as follows. The structure of $\mathcal{T}_0$ is identical to T. For each leaf node in T that contains the singleton element, say a_u, we put the element $a_{\varphi^{-1}(u)}$ in the corresponding leaf node of $\mathcal{T}_0$. The label of each internal node in $\mathcal{T}$ is set to be the diameter of the set of elements contained in the subtree rooted at it.

For each i, j with $1 \le i < j \le \ell$ or $\ell < i < j \le n$, since $i < j$ implies $a_{\varphi(i)} < a_{\varphi(j)}$ by the definition of φ, we have $a_{\varphi(j)} - a_{\varphi(i)} \ge j - i$. Therefore the label of each internal node in $\mathcal{T}_0$ is no larger than that of the corresponding internal node in T. Furthermore, since $a_{\varphi(\ell)} > a_{\varphi(\ell+1)}$ by assumption, we have $a_\ell - a_1 < a_{\varphi(\ell)} - a_{\varphi(1)}$ and $a_n - a_{\ell+1} < a_{\varphi(n)} - a_{\varphi(\ell+1)}$. Therefore, the labels of the roots of the left-subtree and the right-subtree of $\mathcal{T}_0$ are strictly smaller than the labels of their corresponding nodes in T. Hence we can conclude that $\mathcal{R}(\mathcal{T}) < \mathcal{R}(T)$, which is a contradiction to the optimality of T. □ □

Lemma 5. *Let δ_0 be a constant such that our point set cutting lemma holds, then $\delta_0 \ge 2$.*

Proof of Lemma 5. Consider the set of numbers $\mathcal{A}$. Assume that we cut $\mathcal{A}$ at a point $z \in (a_k, a_{k+1}]$, for some $1 \le k < n$. The left-hand side of the inequality in our cutting lemma is $k\cdot(n-k)\cdot(n-1)$, while the right-hand side is $\sum_{1\le i\le k}\sum_{k<j\le n}(j-i) = \frac{1}{2}kn(n-k)$, where the equality follows from Equation (2) derived in § A.4. Therefore we have

$$\delta_0 \ge \frac{k(n-k)(n-1)}{\frac{1}{2}kn(n-k)} = 2\cdot\frac{n-1}{n},$$

which converges to 2 as n tends to infinity. Since this is true for all k with $1 \le k < n$, this lemma follows. □ □

Corollary 6. *Let $\mathcal{M} = (V, d)$ be a given metric and $\mathcal{D}(\mathcal{M})$ be the set of dominating tree metrics of $\mathcal{M}$. Then*

$$\inf_{(V',d')\in\mathcal{D}(\mathcal{M})} \frac{\sum_{u,v\in V} d'(u,v)}{\sum_{u,v\in V} d(u,v)} \geq 2.$$

Proof of Corollary 6. This corollary follows directly from Lemma 4, Lemma 5, and induction on the size of $\mathcal{A}$. □ □

A.4 Computing the Optimal Cut in Linear Time

In this section, we show how the best cut can be computed efficiently in linear time. Let $\{a_1, a_2, \ldots, a_n\}$, $a_1 \leq a_2 \leq \ldots \leq a_n$, be the given set points. For each k with $1 \leq k < n$, let $LS(k) = \sum_{1\leq i<k} (a_k - a_i)$ and $RS(k) = \sum_{k<i\leq n} (a_i - a_k)$ be the sum of the distances between a_k and the points to the left of a_k and the sum of distances between a_k and the points to the right of a_k, respectively. The first observation is that, for $i \leq i < n$,

$$\mathcal{RC}(i) = (n-i) \cdot LS(i) + i \cdot RS(i). \tag{2}$$

The following lemma shows how these quantities can be computed recursively.

Lemma 11. For $1 \leq k < n-1$, We have

- $LS(k+1) = LS(k) + \sum_{1\leq i\leq k} \ell_k$, and
- $RS(k+1) = RS(k) - \sum_{k<i\leq n} \ell_k$.

Proof of Lemma 11. By definition, we have $LS(k+1) = \sum_{1\leq i<k+1} (\ell_k + a_k - a_i) = LS(k) + \sum_{1\leq i\leq k} \ell_k$, and $RS(k+1) = \sum_{k+1<i\leq n} (a_i - a_k - \ell_k) = RS(k) - \sum_{k<i\leq n} \ell_k$. □ □

By Lemma 11 and (2), we can compute in linear time the values $LS(k), RS(k)$, $\mathcal{RC}(k)$ for all $1 \leq k < n$, and the optimal cut. For any given interval $\mathcal{I} \subseteq [a_1, a_n]$, we can also compute the optimal cut inside $\mathcal{I}$ by the same approach.

B Approximating Euclidean Metrics by Their Spanning Trees

ALGORITHM *Euclidean-Spanning-Tree*($\mathcal{P}$)

Input: A set $\mathcal{P}$ of n points in $\mathcal{R}^d$.
Output: A pair $(\mathcal{T}, r)$, which is a spanning tree $\mathcal{T}$ of $\mathcal{P}$ with root r.

1: **if** $\mathcal{P}$ is a singleton point set containing point p **then**
2: Return $(\mathcal{P}, p)$.
3: **end if**
4: Let $\alpha = \frac{1}{4}$ be a constant.
5: Let k be the index of dimension such that $\mathcal{L}_k(\mathcal{B}(\mathcal{P})) = \mathcal{L}_{max}(\mathcal{B}(\mathcal{P}))$.
6: Let $a_1 \leq a_2 \leq \ldots \leq a_n$ be the coordinates of the projection of $\mathcal{P}$ into k^{th} dimension, labelled in sorted order.
7: $p = \alpha \cdot (a_1 + a_n)$, $q = (1-\alpha) \cdot (a_1 + a_n)$.
8: $(\mathcal{P}_1, \mathcal{P}_2) \longleftarrow$ *1d-cut*$(\{a_1, a_2, \ldots, a_n\}, [p, q])$.
9: $(T_1, r_1) \longleftarrow$ *Euclidean-Spanning-Tree*$(\mathcal{P}_1)$, $(T_2, r_2) \longleftarrow$ *Euclidean-Spanning-Tree*$(\mathcal{P}_2)$.
10: Let $T \longleftarrow T_1 \cup T_2 \cup \{(r_1, r_2)\}$.
11: Return (T, r_1).

Figure 5: Algorithm for computing a spanning tree of low routing cost on Euclidean graphs.

For convenience, let $\mathcal{F}$ be the collection of subsets of $\mathcal{P}$ which have occurred during the recursions. For any $\mathcal{Q} \in \mathcal{F}$, we denote by $\mathcal{T}[\mathcal{Q}]$ the subtree of $\mathcal{T}$ corresponding to $\mathcal{Q}$ and $e(\mathcal{Q})$ the edge connecting

the two rooted subtrees corresponding to the two further partitions of $\mathcal{Q}$. $e(\mathcal{Q})$ is defined to be a dummy self-loop with length zero if $\mathcal{Q}$ is a singleton set. The following lemma provides an upper-bound on the pairwise distances.

Lemma 12. For any $p, q \in \mathcal{P}$, we have $d_{\mathcal{T}}(p,q) \leq \frac{2}{\alpha} d\sqrt{d} \cdot \mathcal{L}_{max}(\mathcal{B}(\mathcal{P}))$.

Proof of Lemma 12. Let $A_1 \supset A_2 \supset \ldots \supset A_a$, $A_i \in \mathcal{F}$ for $1 \leq i \leq a$, be the subsets of $\mathcal{P}$ occurred during the recursions to which p belongs, and $B_1 \supset B_2 \supset \ldots \supset B_b$, $B_j \in \mathcal{F}$ for $1 \leq j \leq b$, be the subsets to which q belongs. Note that $A_1 = B_1 = \mathcal{P}$, $A_a = \{p\}$, and $B_b = \{q\}$. From the construction of $\mathcal{T}$, we have

$$d_{\mathcal{T}}(p,q) \leq d_{\mathcal{T}[A_1]}(p, r_1) + |e(\mathcal{P})| + d_{\mathcal{T}[B_1]}(r_2, q) \leq \sum_{1 \leq i \leq a} |e(A_i)| + |e(\mathcal{P})| + \sum_{1 \leq j \leq b} |e(B_j)|,$$

where r_1 and r_2 are the roots of $\mathcal{T}[A_1]$ and $\mathcal{T}[B_1]$. Since the longest straight-line distance inside a hyper-rectangle is bounded by its longest diagonal, we have $|e(Q)| \leq \sqrt{d}\mathcal{L}_{max}(\mathcal{B}(Q))$ for any subset $Q \in \mathcal{F}$. Furthermore, since we always cut along the longest side of the bounding box, we have $\mathcal{L}_{max}(\mathcal{B}(A_{i+d})) \leq (1-\alpha)\mathcal{L}_{max}(\mathcal{B}(A_i))$ and $\mathcal{L}_{max}(\mathcal{B}(B_{j+d})) \leq (1-\alpha)\mathcal{L}_{max}(\mathcal{B}(B_j))$ for all $1 \leq i \leq a-d$ and $1 \leq j \leq b-d$. Therefore, it follows that

$$\begin{aligned} d_{\mathcal{T}}(p,q) &\leq \sum_{1 \leq i \leq a} \sqrt{d}\mathcal{L}_{max}(\mathcal{B}(A_i)) + \sqrt{d}\mathcal{L}_{max}(\mathcal{B}(\mathcal{P})) + \sum_{1 \leq j \leq b} \sqrt{d}\mathcal{L}_{max}(\mathcal{B}(B_j)) \\ &\leq 2d \cdot \sum_{i \geq 1} \sqrt{d}(1-\alpha)^i \mathcal{L}_{max}(\mathcal{B}(\mathcal{P})) + \sqrt{d}\mathcal{L}_{max}(\mathcal{B}(\mathcal{P})) \\ &\leq \frac{2}{\alpha} d\sqrt{d} \cdot \mathcal{L}_{max}(\mathcal{B}(\mathcal{P})), \end{aligned}$$

where in the second last inequality we collect every d items from the summation of the first inequality and then combine them together into a geometric series. □ □

Lemma 7. *Given a set of real numbers $A = \{a_1, a_2, \ldots, a_n\}$, $a_1 \leq a_2 \leq \ldots \leq a_n$ and an interval $\mathcal{I} = [\ell, r]$ such that $\mathcal{I} \subseteq [a_1, a_n]$, there exists a cutting point $z \in \mathcal{I}$ such that the following holds.*

$$L_A(z) \cdot (n - L_A(z)) \cdot |\mathcal{I}| \leq \delta_0 \cdot \sum_{1 \leq i \leq L_A(z)} \sum_{L_A(z) < j \leq n} (a_j - a_i),$$

where $L_A(z) = |\{a \in A : a < z\}|$ is the number of elements in A that are smaller than z and $\delta_0 \leq \frac{210}{59}$ is a constant.

Proof of Lemma 7. We say that an interval degenerates if it has length zero. First we argue that, if there are degenerating intervals at a_1, then it is always worse to cut at those degenerating intervals. Let k, $1 \leq k \leq n$, be the largest index such that $a_1 = a_2 = \ldots = a_k$. Observe that, for any i, j with $1 \leq i, j \leq k$, we have $\mathcal{RC}(i) = \frac{i}{j} \cdot \mathcal{RC}(j)$. On the other hand, for $1 \leq i < k$ and $1 \leq j \leq k - i$, we have

$$\frac{(i+j)(n-i-j)}{i(n-i)} = \frac{i(n-i) + j(n-2i-j)}{i(n-i)} \leq \frac{i+j}{i} = \frac{\mathcal{RC}(i+j)}{\mathcal{RC}(i)},$$

which implies that $\frac{(i+j)(n-i-j)}{\mathcal{RC}(i+j)} \leq \frac{i(n-i)}{\mathcal{RC}(i)}$ and therefore cutting at $(a_k, a_{k+1}]$ is always better than cutting at degenerating intervals at a_1. Similarly, we can argue that, it is always worse to cut at the degenerating intervals at a_n, if there is any.

Now we argue that there will be a feasible cut satisfying the criterion. According to the given interval $\mathcal{I} = [a, b]$ and the point set A, we create a new point set $B = \{b_1, b_2, \ldots, b_n\}$ as follows.

$$\text{For } 1 \leq i \leq n, \quad b_i = \begin{cases} \ell & \text{if } a_i < \ell, \\ a_i & \text{if } \ell \leq a_i \leq r, \\ r & \text{otherwise.} \end{cases}$$

Let z be the best cut of B in $\mathcal{I}$. By the above argument, we have $\ell < z < r$ and therefore $L_A(z) = L_B(z)$. By Lemma 3, we have $L_B(z) \cdot (n - L_B(z)) \cdot |\mathcal{I}| \leq \frac{210}{59} \sum_{b_i < z \leq b_j} (b_j - b_i)$. According to our setting, we have $(b_j - b_i) \leq (a_j - a_i)$ for all $1 \leq i < j \leq n$. Therefore $L_A(z) \cdot (n - L_A(z)) \cdot |\mathcal{I}| \leq \frac{210}{59} \sum_{1 \leq i \leq L_A(z)} \sum_{L_A(z) < j \leq n} (a_j - a_i)$ as claimed. □ □

Theorem 8. *Given a set of points $\mathcal{P}$ in $\mathcal{R}^d$, Algorithm* Euclidean-Spanning-Tree *computes a spanning tree $\mathcal{T}$ of $\mathcal{P}$ such that the distance-weighted average stretch of $\mathcal{T}$ with respect to $\mathcal{P}$ is at most $16\delta_0 \cdot d\sqrt{d}$, where $\delta_0 \leq \frac{210}{59}$ is the constant in our point set cutting lemma.*

Proof of Theorem 8. If $|\mathcal{P}| = 1$, then this theorem holds trivially. Otherwise, by Lemma 12, Lemma 7, and the fact that the length of the restricted interval is $(1 - 2\alpha) \cdot \mathcal{L}_{max}(\mathcal{B}(\mathcal{P}))$, we have

$$\mathcal{R}_{\mathcal{T}}(\mathcal{P}_1, \mathcal{P}_2) \leq |\mathcal{P}_1| \cdot |\mathcal{P}_2| \cdot \frac{2}{\alpha} d\sqrt{d} \cdot \mathcal{L}_{max}(\mathcal{B}(\mathcal{P})) \leq \frac{2\delta_0}{\alpha(1-2\alpha)} d\sqrt{d}\mathcal{R}(\mathcal{P}_1, \mathcal{P}_2).$$

This holds for all recursions. Choose α to be $\frac{1}{4}$ and this theorem follows directly by induction on the depth of recursion. □ □